

Instructions for authors, subscriptions and further details:

<http://ijep.hipatiapress.com>

How Much Guidance Do Students Need? An Intervention Study on Kindergarten Mathematics with Manipulatives

Erin Horan¹, Martha Carr²

- 1) Association of American Colleges and Universities, USA
- 2) University of Georgia, USA

Date of publication: October 24th, 2018
Edition period: October 2018 – February 2019

To cite this article: Horan, E., & Carr, M. (2018). How Much Guidance Do Students Need? An Intervention Study on Kindergarten Mathematics with Manipulatives. *International Journal of Educational Psychology*, 7(3), 286-316. doi: [10.17583/ijep.2018.3672](https://doi.org/10.17583/ijep.2018.3672)

To link this article: <http://dx.doi.org/10.17583/ijep.2018.3672>

PLEASE SCROLL DOWN FOR ARTICLE

The terms and conditions of use are related to the Open Journal System and to [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/) (CC-BY).

How Much Guidance Do Students Need? An Intervention Study on Kindergarten Mathematics with Manipulatives

Erin Horan
*Association of American Colleges
and Universities*

Martha Carr
University of Georgia

Abstract

Research has shown that the efficacy of learning with manipulatives (e.g., fingers, blocks, or coins) is affected by multiple variables, including the amount of guidance teachers provide during learning. However, there is no consensus on how much guidance is necessary when learning with manipulatives. The goal of this study was to examine the optimal level of guidance during instruction with manipulatives. The focus was on the timing and level of guidance. The researcher taught students a lesson on counting from one to 10 with pennies and nickel strips. Kindergarten students were taught over five consecutive days in one of four conditions: high guidance, low guidance, high guidance that transitioned to low guidance, and low guidance that transitioned to high guidance. Results showed no difference in learning across the conditions. These results provide valuable information to teachers on the areas of mathematics that do not require the effort of high guidance.

Keywords: mathematics education, elementary school, manipulatives, guidance



¿Cuánta Guía Necesitan los Estudiantes? Una Intervención de Enseñanza de Matemáticas en Estudiantes de Educación Infantil con Manipulativos

Erin Horan
*Association of American Colleges
and Universities*

Martha Carr
University of Georgia

Resumen

Múltiples estudios han demostrado que la eficacia del aprendizaje con medios manipulativos (por ejemplo, dedos, bloques, o monedas) está relacionada con múltiples variables, incluyendo la guía que proveen las maestras y los maestros durante el aprendizaje. Sin embargo, no existe consenso sobre cuánta guía es necesaria durante el aprendizaje con medios manipulativos. La meta de este estudio fue examinar el nivel óptimo de guía necesaria durante el aprendizaje con medios manipulativos. El estudio se enfocó en los momentos y el nivel de orientación. La investigadora enseñó a estudiantes una lección sobre cómo contar del 1 al 10 haciendo uso de monedas de un centavo y tiras de papel con cinco monedas de un centavo dibujadas a un lado y una moneda de cinco centavos al otro. Durante cinco días consecutivos, la lección se impartió a estudiantes de escuela infantil en una de las siguientes cuatro condiciones: nivel de guía alto, nivel de guía bajo, nivel de guía alto con transición a un nivel de guía bajo, y nivel de guía bajo con transición a un nivel de guía alto. Los resultados no demostraron diferencias en aprendizaje entre las cuatro condiciones. Estos resultados proveen información valiosa para maestras y maestros en las áreas de matemáticas que no requieren el esfuerzo de un nivel alto de guía.

Palabras clave: educación en matemáticas, escuela infantil, manipulativos, guía



The utility of manipulatives to support learning has been widely accepted and recommended (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). However, investigations by Carbonneau and her colleagues (Carbonneau & Marley, 2015; Carbonneau, Marley, & Selig, 2013) have shown the efficacy of learning with manipulatives is not consistent and depends on many variables related to the instruction including the level of guidance (e.g., high guidance or low guidance) and the student's prior knowledge. There is evidence that at least some guidance during mathematics instruction is necessary for optimal learning but the literature is unclear as to when teachers should provide guidance and when they should allow students to practice alone without teacher help. In the study described below, we examined students' learning with manipulatives (pennies and nickel strips) with varying levels of guidance. We implemented an experiment in which the amount and timing of guidance with manipulatives was tested using four conditions.

Research on Manipulatives and Guidance

Manipulatives refer to any concrete materials, objects, or drawings used during instruction to support students' learning of number and operations. Manipulatives can be simple, such as counting on fingers or unit blocks, or complex, such as using base ten sticks and blocks. In elementary school mathematics classrooms, students learn to count using individual manipulatives to determine "how many" (National Research Council, 2009). Later, students move on to complex manipulatives that represent values of the base-ten system. In elementary school, manipulatives are incorporated into mathematics curricula to aid students' mathematics reasoning and problem solving skills (e.g., Expressions, Investigations, Saxon).

Sowell (1989) conducted a meta-analysis on the effectiveness of using manipulatives during mathematics instruction and found that using manipulatives was better than not using manipulatives. Younger students, especially, benefitted from using manipulatives as they provide concrete objects to students who may not yet be able to think abstractly (DeLoache, 2000; Uttal, O'Doherty, Newland, Hand, & DeLoache, 2009). Carbonneau, Marley, and Selig (2013) followed up on this research and conducted a meta-analysis of 55 studies that explored the efficacy of teaching with complex manipulatives and found that teaching with manipulatives compared to teaching with abstract symbols showed small to medium sized effects on

student learning. The research has shown that manipulatives can aid learning, but there are certain variables (i.e., guidance and prior knowledge) that can mitigate their helpfulness.

The term guidance has been used to describe many types of instructional formats. For example, some research has used the term guidance to describe student-teacher interactions that occur during the learning process (e.g., Terwel, van Oers, van Dijk, & van den Eeden, 2009; Mayer, 2004). Other research uses guidance to describe other aspects of instruction, such as guiding students by providing worked examples, formula sheets, or systematically ordering problems to lead students to insightful learning experiences (e.g., Baroody, Purpura, Eiland, & Reid, 2015; Chen, Kalyuga, & Sweller, 2015). Horan (2017) discussed how guidance is used to describe a variety of instructional components as well as the issues that stem from the lack of a clear definition of guidance. For the purposes of this study, guidance is defined as the interaction between a teacher and students, specifically, the quantity and quality of teachers' responsiveness to students' questions and concerns, and teachers' tendency to promote reflection and critical thought with questions and comments. Examples of high quality interaction include a teacher monitoring student response during problem solving and providing assistance as needed, teachers providing feedback and responding to questions from students, students responding verbally to questions from teachers, and teachers creating opportunities for reflection based on students' performance and needs. In contrast, simply providing performance feedback (i.e., correct or incorrect) that is not responsive to students' needs would be considered low guidance.

Overall, prior research shows support for implementing instruction with manipulatives in the mathematics classroom. However, the research on guidance, especially guidance with manipulatives, is less clear. Further, understanding how prior knowledge impacts guidance with manipulatives adds another variable to investigate.

Guidance and Prior Knowledge

Regarding the effectiveness of guidance when using manipulatives, Laski, Jordan, Daoust, and Murray (2015) summarized general findings on young children's learning with mathematics manipulatives and recommended the use of explicit guidance that relates the concrete manipulatives to the abstract

numbers they represent. Providing consistent guidance was found to allow students to devote working memory to understanding the content of the mathematics lesson rather than other, extraneous content. In their meta-analysis, Carbonneau, Marley, and Selig (2013) found that high guidance instruction was associated with higher retention and problem solving performance, while low guidance instruction was associated with higher transfer performance when using manipulatives. Carbonneau et al. (2013) also investigated the impact of age on learning and found that students age 3-6 (preoperational age) struggled more when learning with manipulatives compared with students in the concrete operational age group (7-11) or formal operational age group (12 and older). The authors attributed this finding to young students' tendency to struggle with understanding that objects can represent larger mathematical concepts.

While there is clear evidence high guidance is useful for learning, there is some evidence that there are benefits to implementing lessons with low guidance. Therefore, we are interested in understanding the benefits of high versus low guidance, as well as instruction that transitions the level of guidance during learning (e.g., high to low guidance, low to high guidance). We also look to the research on prior knowledge for learning mathematics, to further understand how prior knowledge may determine the usefulness of high or low guidance on learning.

Support for high guidance. Support for high guidance instruction comes from researchers and theorists who argued that without teacher guidance, students left to their own devices will not learn concepts or, worse, learn the wrong concepts (Rogoff, 1990; Cobb, 1995). Social constructivist theorists posit that high guidance during learning with manipulatives is essential because the manipulatives are culturally-specific, external representations that allow children to count before having an internal representation of number. In order to support the eventual development of an internal representation of numbers, students need guidance to be able to recognize what the concrete manipulatives represent (Bruner, 1966; Vygotsky, 1978).

More recent, empirical research supports implementing high guidance during learning, explaining that exploration without the guidance of an instructor can result in students never interacting with the content to be learned (Mayer, 2004). For example, Terwel, van Oers, van Dijk, and van den Eeden (2009) compared the impact of two problem-solving lessons on student learning of percentages and graphs. In the high guidance condition, fifth grade students

were taught through the process of guided co-construction; students and teachers created representations of the percentages through teacher-initiated, guided discussions. In the low guidance condition students were provided with ready-made, completed representations and were not engaged in discussion with the teacher. Controlling for pretests scores, children in the high guidance condition performed better on a posttest and transfer test. This provided support for guided, interactive teaching when students are learning problem solving strategies for percentages and graphs.

Fisher, Hirsh-Pasek, Newcombe, and Golinkoff (2013) described guided instruction as a collaborative construction by students and teachers. In their study, Fisher et al. (2013) taught preschool students properties of shapes in three conditions: free play in which student activity was self-directed with no goals for learning (i.e., low guidance), a guided play condition described as discovery learning with the presence of an active teacher participant, and an instruction condition in which the student observed the instructor talking through the material. The authors found that students in the guided play condition showed improved understanding of shapes over the other two conditions, and those improvements were still observed one week later. They found that for understanding properties of shapes, high guidance, even when scripted, was better than instruction that involved the student passively listening to the teacher or playing alone without any guidance.

Carbonneau and Marley (2015) also worked with preschool students to compare the impact of different levels of guidance. The study investigated the impact of guidance on students' conceptual and procedural knowledge on a quantity discrimination task (which side has more) using manipulatives. In their study, the researcher would make two piles of objects and the child would have a crocodile mouth with instructions that the crocodile should eat the bigger number. After making the piles the researcher would ask, "Which one should the crocodile eat?" In one condition, which the authors labeled high guidance, after the child pointed to the pile the crocodile should eat, the researcher would then read the number sentence represented by the piles and crocodile and correct the child if necessary. In the low guidance condition, the researcher prompted the student to read the number sentence. Carbonneau and Marley (2015) found that students who heard the teacher repeat their explanations and were corrected on their errors improved their conceptual and procedural knowledge more than students who only received prompts to recite the number sentence on their own.

Support for low guidance. While the importance of high teacher guidance is evidenced by prior research, others research shows that high guidance does not always lead to improved performance over low guidance. For example, Sengupta-Irving and Enyedy (2014) compared a guided condition, where the teacher led fifth grade students through the problem solving process via interactive discussion to an unguided, open approach, where students completed the problem without any assistance from the teacher. They found no group differences in learning outcomes between conditions on data analysis and probability.

Further, there are situations where not only is high guidance not any better than low guidance, but low guidance is more effective than high guidance. One reason for this it is important to provide learners with time for their own exploration (e.g., Bruner, 1961; Schwartz, 1992). Low guidance instruction gives learners the opportunity to formulate and understand mathematical concepts on their own, which is important for deeper learning of mathematics knowledge (Piaget, 1977; Fuson, 2009). Low guidance can also avoid the effects of overwhelming students' working memory with too many questions or comments from a teacher (Kroesbergen and Van Luit, 2005). It is important to note that pure discovery learning, where students are left with no guidance or instruction and only materials, has not been found to help students learn; instead, researchers advocate for learning that incorporates some outside assistance in the form of feedback on steps the student is taking or outcome feedback on their answers (Alfieri et al., 2011).

Looking to empirical support for low guidance instruction, Kroesbergen and Van Luit (2005) found low guidance was better than high guidance instruction for students with mild intellectual disabilities who were learning multiplication solution procedures. Kroesbergen and Van Luit (2002) likewise found that students in special education classes benefitted more from low guidance than high guidance, however they found that low performing students not identified as having a learning disability benefitted more from high guidance. These findings indicated that students with learning disabilities may have characteristics that differentiate the impact of guidance on learning. While the current study did not implement research with students with disabilities, we chose these examples to highlight the many variables to consider when researching guidance in mathematics instruction, especially given the limited literature on guidance during mathematics instruction with manipulatives. Additional research, which shows support for low guidance

instruction depending on students' prior knowledge, is discussed later with the effects of prior knowledge.

Support for transitioning guidance. Another approach to implementing guidance involves starting with high guidance and then transitioning to low guidance as students gain skill and fluency. This format of ordering guidance was studied by Fuchs et al. (2003) who investigated whether initial high guidance instruction followed by exploratory problem solving is superior to exploration followed by guided instruction. Fuchs et al. (2003) found that problem solving improved for students who had high guidance instruction followed by low guidance problem solving with fully worked examples compared to a high guidance, instruction-only condition. However, high guidance instruction followed by low guidance problem solving with partially worked examples, rather than fully worked examples, was not better than high guidance, instruction-only. These findings showed that the optimal level and timing of guidance may depend on multiple variables, such as the age of students, mathematical topic, and structure and content of the instruction or problems. This points to the need for further research on transitioning levels of guidance.

The role of prior knowledge. Cognitive load theory stipulates that students with less prior knowledge need more guidance so as not to exceed their cognitive load. Students with more domain specific knowledge will not need as much guidance because the information is stored in long term memory (Sweller, Ayres, & Kalyuga, 2011). Guidance should be given to support the acquisition of new knowledge, and not to focus on information that has already been learned because this could confuse the students if conflicting information is given (Kalyuga, 2007). This means teachers need to monitor the amount of guidance to give based on students' prior knowledge and experience with a topic.

Fyfe and Rittle-Johnson (2016) investigated the impact of computer feedback on second grade students' learning of equivalency problems. There were three conditions within computer-based problem solving: no-feedback; immediate accuracy feedback after each problem; and summative, accuracy feedback after all 12 problems were solved. Within each of these three conditions students were grouped as having high or low prior knowledge. The impact of feedback differed as a function of prior knowledge. Students with lower prior knowledge, performed better in the feedback conditions than no feedback conditions on solving equivalency problems. For students with

higher prior knowledge, all conditions resulted in improvement on solving equivalency problems.

Jitendra et al. (2013) found a different effect of prior knowledge on learning with high and low guidance. They compared a high guidance condition that utilized schema-based instruction to a low guidance, business-as-usual group. The high guidance condition involved a curriculum in which the teacher prompted students to use think-alouds to encourage monitoring and reflection during problem solving. The low guidance condition involved a school-provided, inquiry-based curriculum, in which students worked alone to develop multiple solutions for an ordered set of problems presented on worksheets. Surprisingly, students with higher pretest scores (high prior knowledge) were found to perform significantly better with the high guidance, schema-based curriculum whereas students with lower pretest scores performed better with the low guidance curriculum. Tournaki (2003) compared performance on mathematics addition tasks for second grade students, half of which were general education students and half of which were students with learning disabilities. For students with learning disabilities, significant improvements from pretest to posttest were only found for students in the high guidance instruction group. General education students improved in both the low and high guidance groups. For both the general education students and the students with learning disabilities significant improvements on the transfer task were only found for students in the high guidance condition.

The results of the studies discussed do not paint a clear picture of the role of prior knowledge. Carbonneau, Marley, and Selig (2013) found that high guidance interventions with manipulatives produced better retention than low guidance interventions, but low guidance interventions produced better transfer than high guidance interventions. Another alternative is to include both high and low guidance in the instruction and determine the optimal sequence of guidance (e.g., Darch, Carnine, & Gersten, 1984; Fyfe, Rittle-Johnson, & DeCaro, 2012). Even further, the optimal level or sequence of guidance may also be influenced by students' prior knowledge (e.g., Jitendra et al., 2013; Tournaki, 2003).

Current Study

In the current study we compared student performance on measures of mathematics achievement after one of four five-day treatments that differed in the amount and/or timing of guidance. In the high guidance condition, students were taught with consistent high guidance for all five days. In the low guidance condition, students were taught with low guidance for all five days. In the high to low guidance condition, students were taught with high guidance for the first two days, low guidance for the last two days, with the third day utilized as a transition day where the researcher limited the guidance but did not eliminate it until day 4. In the low to high guidance condition, students were taught with low guidance for the first two days, high guidance for the last two days, with the third day utilized as a transition day where the researcher added some high guidance questions and comments. Our study specifically investigated four questions:

1. How does student performance on measures of mathematics differ based on teacher guidance when using manipulatives?
2. How does teacher guidance impact kindergarten student performance on a transfer task?
3. How does teacher guidance impact kindergarten student performance on a measure of number sense?
4. To what extent does the effect of teacher guidance differ based on kindergarten student prior knowledge (initial skill)?

Carbonneau et al. (2013) found high guidance was optimal for improving student performance on the task being taught. Research has also shown support for low guidance at some point during instruction, but it is not clear if high guidance should be faded out or if it should come after low guidance instruction (e.g., Kroesbergen and Van Luit, 2002, 2005). Therefore, we predicted that one of the transitioning conditions (high to low or low to high) would be best for impacting student performance on the counting task being taught.

Carbonneau et al. (2013) also found that studies that implemented low guidance interventions with manipulatives had higher effect sizes for transfer than the studies that implemented high guidance interventions with manipulatives. On the other hand, students in low guidance instruction group may not learn at all, and may need guidance from the teacher to learn not just

the material, but enough to be able to transfer to another task. Prior studies have found lower achieving students need more guidance to understand the content in order to transfer knowledge (Tournaki, 2003). We were interested in understanding transfer effects because we predicted the transitioning conditions would lead to students learning the counting task at hand, but we wondered if the benefits would also transfer to other situations as other studies have investigated. First, we predicted the consistently low guidance condition would not be the optimal condition for transfer because not all students would be able to learn completely on their own without any guidance. We predicted that the low to high and high to low guidance conditions would lead to better transfer because students would have the opportunity to make meaningful connections on their own. This was also our prediction for posttest performance on the Test of Early Numeracy (TEN) as the TEN can be considered far transfer and the same issues and predictions held for the impact of the different conditions on the TEN.

Based on the review by Kalyuga (2007) on the expertise reversal effect, we hypothesized there would be an interaction effect with prior knowledge. Students with low prior knowledge would perform best with consistent high guidance or high to low guidance to learn with manipulatives. If students are not given enough guidance to start with they may learn information incorrectly or may not know where to begin when exploring with manipulatives alone. Students with high prior knowledge may need consistent low guidance or low to high guidance to learn with manipulatives. These students need time to explore alone and already have enough prior knowledge to do this effectively. Starting with high guidance may confuse students with high prior knowledge.

Method

Participants

Consent forms were distributed to kindergarten students at four elementary schools from a southeastern school district. Students at this school district are comprised of 61% white, 17% Hispanic, 13% black, 5% multi-racial, and 4% Asian. One hundred sixty-seven students consented to participate. Of those, one student was absent during the week of the intervention and one student with special needs could not complete the measures for testing so the final sample was 165 (99 males, 66 females).

The sample was comprised of students who were 71.5% white, 13.3% Hispanic, 12.1% black, and 3% Asian. At the start of the study in fall 2015, the average age of students was 5.56 years, $SD=0.36$. Students came from three different schools. Seven classes participated from the first school, eight classes participated from the second school, and five classes participated from the third school.

It should be noted that several teachers requested to send only students who might benefit from the intervention so as not to have too many students missing class time. As such, some teachers only sent consent forms home with students of their choosing. While this is beneficial for the purposes of understanding how guidance impacts students not performing as well as their peers in mathematics, this does limit the generalizability of our study. Further, as we do not have an accurate count of how many students were originally recruited and asked to participate, we cannot determine the percent of recruited participants who consented to participate.

Design

Students were randomly assigned to one of the four conditions. The lessons were taught by the first author (referred to as researcher) who designed the study. In the academic year prior to the study, the researcher implemented a pilot study with pre-school students to ensure the feasibility of conducting this study with groups of young students. The lessons took place in conference rooms as available, which are typically limited to five to seven seats. Throughout the school day, the researcher pulled students from class in groups of five to seven. Students from different classes would be combined to form groups and the groups could change from day to day. For example, if one class was in the middle of an important lesson, the researcher would go to other classes to pull other students to form the full group. Teaching took place for six to nine minutes per day for five days.

Materials and procedure

All students were assessed at pretest, posttest, and delayed posttest on a counting with manipulatives task, the Test of Early Numeracy (TEN), and a transfer task. All three tests were given at all three time points. The pretest was administered the week before the intervention took place. The posttest was administered the week after the intervention and a delayed posttest was

administered two weeks after the intervention. All pretest, posttest, and delayed posttest measures were individually administered in a quiet area. Administration took approximately 10-15 minutes per student, per testing occasion.

Counting task. The counting tasks were designed to assess student ability to count manipulatives. The counting tasks utilized ten boards which are used to teach kindergarten students the order of numbers as part of the Math Expressions curriculum. It should be noted that Math Expressions was not the curriculum used by the school district. The counting task designed for the pretest was different from the counting task designed for the posttests because students had not yet been introduced to the nickel strips at the time of the pretest and the researcher did not want to provide instruction on the nickel strips until the time of the intervention. For the pretest, students were given a board with the numbers one through ten at the top of the board. Below each number was a column for the student to place pennies to show the value of the number (see Figure 1). The researcher asked students, “Place the number of pennies that are written at the top of the column”. The students used pennies to show the numbers given. For the pretest the researcher asked students to place the correct number of pennies under the columns five, eight, three, one, and six. The nickel strips were not used for the pretest. As this task had students fill in five total columns the scores for the pretest counting task could range from 0-5.

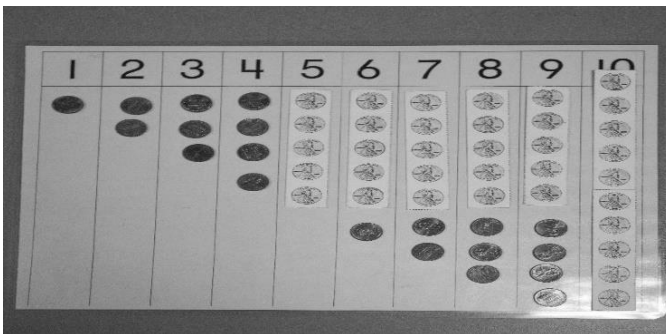


Figure 1. Intervention and the intervention specific task. One through ten board from Math Expressions. The instructions were: "You have a board, pennies and nickel strips. You are going to make the numbers one through ten on your boards using pennies and nickel strips. You can use pennies to make all of the numbers. You can

also use nickel strips. Each nickel strip stands for five pennies (show nickel strip which has pictures of five pennies on it). You can use one nickel strip to take the place of five pennies. The number in the first column is one, can you make the number one with pennies?"

The posttest and delayed posttest had students place all pennies and nickel strips under all columns from one to ten, as shown in [Figure 1](#). This measure was scored as either correct or incorrect which resulted in two categories; 0=not correct; 1=correct. To be scored correctly students needed to place the pennies and nickel strips correctly. The possible range of scores for this task was 0-1. Cronbach's alpha for the pretest, posttest, and delayed posttest for the consistent low guidance condition found counting tasks to be reliable (3 tests; $\alpha = .529$).

Number sense measure. Number sense was measured with the Test of Early Numeracy (TEN). The TEN ([Clarke & Shinn, 2004](#)) is individually administered and includes four measures; each measure lasts for one minute for a total of about five minutes per student. The four measures on the TEN are oral counting (possible scores 0-100), number identification (possible scores 0-56), quantity discrimination (possible scores 0-28), and missing number (possible scores 0-21). The oral counting measure has students count as high as they can for one minute. The number identification measure has students identify numbers between 1 and 10 for kindergarteners. The quantity discrimination measure has students identify the larger of two numbers between 1 and 10 for kindergarteners. The missing number measure has students identify the missing number for a set of three numbers with two numbers given. Rather than one, summative score, the TEN yields four separate scores for number sense, which were analyzed individually.

The TEN has been shown to be a valid measure of number sense for kindergarten and first grade students. [Clarke and Shinn \(2004\)](#) found the TEN was correlated with the Woodcock-Johnson Applied Problems ([Woodcock & Johnson, 1989](#)) subtest for first grade students, which measures mathematics achievement based on mathematics operations problems and applied mathematics problem. [Martinez, Missall, Graney, Aricak, and Clarke \(2009\)](#) found that the TEN was correlated with Stanford 10 Achievement Test ([Harcourt Assessment Inc., 2002](#)), which measures if students are meeting standards for reading, mathematics, and language. Alternate form reliability

was measured by testing students with an alternate form of all subtests except for the oral counting measure because there is no alternate form for counting as high as you can (Clarke & Shinn, 2004). Reliabilities for the TEN were measured as .93 for oral counting, .93 for number identification, .92 for quantity discrimination, and .78 for missing number (Clarke & Shinn, 2004). Salvia and Ysseldyke (2001) assigned a reliability of .90 or greater for making educational decisions about individual students, .80 or greater for making screening decisions about individual students, and .60 or greater for making educational decisions about groups of students. According to these guidelines all measures of the TEN can be used to make educational decisions about individuals except the missing number measure, but .78 is still a moderately high reliability.

Transfer task. Transfer was assessed with a task that required students to count on from five. Students were shown a number between six and ten and five circles. Sample transfer problems are shown in Figure 2. Students were given the following instructions: “Do you see that we have 1,2,3,4,5 circles? Can you draw more circles so we have X circles in the box?” This task was scored as zero correct, one correct, or both correct. Cronbach’s alpha for the pretest, posttest, and delayed posttest for the consistent low guidance condition found the transfer task to be reliable (3 tests; $\alpha = .667$). We chose to report reliability for the consistent low guidance condition as this condition was essentially a control condition; students were not given any guidance during the intervention.

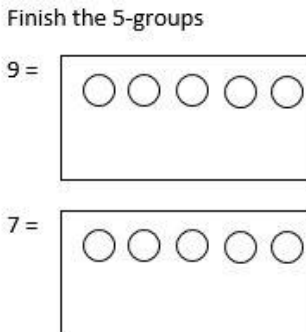


Figure 2. Transfer worksheet for pre-test post-test, and delayed post-test.

Intervention

As described in more detail below, the consistent high guidance group implemented only the high guidance lesson throughout the entire week and the consistent low guidance group implemented only the low guidance lesson throughout the entire week of the intervention. The high to low guidance group began the week with high guidance lessons then shifted to low guidance lessons. The low to high guidance group began the week with low guidance lessons and then shifted to high guidance lessons. To assure fidelity of the high and low guidance modifications, all lessons were recorded and coded as described below.

The four conditions utilized ten boards which are used to teach kindergarten students the order of numbers and are a part of the Math Expressions curriculum (see Figure 1). For this task, students make the numbers one through ten by using pennies and nickel strips. Nickel strips are white pieces of paper that fit perfectly under five pennies. The students first counted out the number of pennies requested then added a nickel strip under sets of five pennies. For example, if the number eight was counted the student would count out eight pennies, then replace five of those pennies with a nickel strip.

High guidance modification. Three of the four conditions include lessons that have high guidance. A high guidance lesson is defined as the teacher asking many questions during learning. A list of possible questions is included in Table 1. The teacher was not required to use every question on this list nor was the list an exhaustive list of questions asked. The high guidance instruction used these questions to increase student learning and understanding. The teacher also provided elaborate feedback about performance during the lesson (not just right or wrong but why), helped students if they needed help, and answered students' questions.

Table 1

Sample High guidance questions.

“Why did you (not) use a nickel strip in this column?”

“How is this column different from the last column?”

“How is the 8 column the same as the 3 column? How is it different?”

“Can we use a nickel strip in this column? Why (not)?”

“Can you count the pennies to check your answer?”

“How many more pennies would we need for a nickel strip?”

“How many more pennies would we need for another nickel strip?”

“How many more pennies would we need for 5?”

“How many more pennies would we need for 10?”

Low guidance modification. Three of the four conditions include lessons that have low guidance. Per the definition of low guidance for this paper the teacher could provide feedback to students in the form of “yes” or “no” but provided no further information. In the case of the activity to learn the numbers one to ten, low guidance included instructions to make the numbers one to ten and corrective feedback, but did not include any back and forth questioning. In addition to the instructions given in [Figure 1](#), students were provided the following instruction once they reached the five column: “When you reach the number five on the number board you take away the five pennies and use a nickel strip instead.” For the numbers six through 10 these instructions were repeated. Questions to keep the students on task could be asked, but questions about the content (e.g., “which number is bigger?”) were not.

Transitioning conditions. There were two transitioning conditions; high to low guidance and low to high guidance. For the first two days students were taught with either high or low guidance. Day three was a transition day where the level of guidance started to taper off so that high guidance was tapered to low guidance or increased so that low guidance increased to high guidance. On days four and five students were taught with the second type of guidance so that students who were given low guidance on days one and two were given

high guidance and students who were given high guidance on days one and two were given low guidance.

Fidelity. To ensure fidelity of the high and low guidance modification all lessons were audio recorded and coded. Each lesson was rated as high guidance or low guidance based on the number of questions asked by the researcher to students; less than five indicated low guidance, more than five indicated high guidance. Five was chosen as the cutoff to allow room for the low guidance conditions to include minimal questioning such as to keep students on task, as completely cutting out questions is not realistic or practical in everyday teaching. The researcher and a trained independent rater (a graduate student) coded 20% (31) of the sessions. Interrater reliability between the researcher and the independent coder was established as 96.8%.

Results

Performance on Counting with Manipulatives

Means and standard deviations of performance on the pretest, posttest, and delayed posttest counting tasks for students in each condition are shown in Table 2. Analysis of variance showed that pretest scores on the content task did not significantly differ across conditions, $F(3,164) = 1.506$, $p = .225$. Posttest scores on the content task controlling for pretest scores were not significantly different across conditions, $F(3,160) = 0.735$, $p = .532$. Delayed posttest scores on the content task controlling for pretest scores were not significantly different across conditions, $F(3,160) = 1.128$, $p = .339$.

Table 2

Mean Performance on Content Tasks and Standard Deviation.

	H-H	H-L	L-H	L-L	Full sample
Pretest	4.73 (.554)	4.49 (1.00)	4.54 (.745)	4.72 (.701)	4.62 (.769)
Posttest	0.68 (.474)	0.63 (.488)	0.51 (.506)	0.63 (.489)	0.61 (.489)
Delayed Posttest	0.70 (.464)	0.63 (.488)	0.51 (.506)	0.70 (.465)	0.64 (.483)
<i>N</i>	40	41	41	43	165

Note: Pretests were scored as 0-5. Posttests and delayed posttests were scored as 0 (incorrect) or 1 (correct).

Subsequent analyses were performed after removing students who scored with the highest score on the pretest (five out of five) to account for ceiling effects. Means and standard deviations of performance on the pretest, posttest, and delayed posttest counting tasks for students in each condition are shown in Table 3. Again, posttest scores on the content task controlling for pretest scores were not significantly different across conditions, $F(3,38) = 0.040$, $p = .989$. Delayed posttest scores on the content task controlling for pretest scores were not significantly different across conditions, $F(3,38) = 0.126$, $p = .944$.

Table 3

Performance on Content Tasks with Highest Performers Removed.

	H-H	H-L	L-H	L-L	Full sample
Pretest	3.78 (.441)	3.38 (1.19)	3.64 (.633)	3.29 (.756)	3.53 (.827)
Posttest	0.56 (.527)	0.54 (.519)	0.57 (.514)	0.57 (.535)	0.56 (.502)
Delayed Posttest	0.56 (.527)	0.54 (.519)	0.50 (.519)	0.57 (.535)	0.53 (.505)
<i>N</i>	9	13	14	7	43

Note: Pretests were scored as 0-5. Posttests and delayed posttests were scored as 0 (incorrect) or 1 (correct).

Performance on Transfer

The transfer pretest, posttest, and delayed posttest were two item tasks scored from 0-2. Means and standard deviations of performance on the pretest, posttest, and delayed posttest transfer tasks for students in each condition are shown in Table 4. Analysis of variance showed that pretest scores on the transfer task did not significantly differ across conditions, $F(3,164) = 0.283$ $p = .838$. Posttest scores on the transfer task controlling for pretest scores were not significantly different across conditions, $F(3,160) = 0.544$ $p = .653$. Delayed posttest scores on the transfer task controlling for pretest scores were not significantly different across conditions, $F(3,160) = 0.618$ $p = .604$. Subsequent analyses were performed after removing students who scored with the highest score on the pretest (five out of five) to account for ceiling effects but these findings were not significant.

Table 4

Mean Performance on Transfer Tasks and Standard Deviation Controlling for Pretest.

	H-H	H-L	L-H	L-L	Full sample
Pretest	1.53 (.784)	1.46 (.745)	1.37 (.767)	1.47 (.855)	1.45 (.784)
Posttest	1.63 (.667)	1.61 (.628)	1.44 (.776)	1.63 (.618)	1.58 (.673)
Delayed Posttest	1.68 (.616)	1.59 (.741)	1.71 (.602)	1.60 (.660)	1.64 (.653)
<i>N</i>	40	41	41	43	165

Note: Pretest, posttest, and delayed posttest were scored from 0-2.

Performance on TEN

A three stage hierarchical linear regression was conducted with each posttest TEN score as the dependent variable. Pretest TEN score was entered at stage one of the regression to control for prior knowledge. The four conditions were dummy coded into three variables and were included at stage two. Interactions between the conditions and pretest scores were included at stage three. Performance on each component of the TEN (i.e., counting, missing number, number identification, and quantity discrimination) was analyzed separately.

The hierarchical multiple regression revealed that at stage one, pretest scores on each component of the TEN contributed significantly to the regression model. Beyond stage one, only the model for the counting component of the TEN showed significant contributions by other variables (see Table 5). For the counting component, pretest scores contributed significantly to the regression model, $F(1,163) = 326.0, p < .001$ and accounted for 66.7% of the variation in posttest scores. Introducing the experimental conditions explained an additional .3% of the variation in posttest scores and this change in R^2 was significant, $F(4,160) = 81.1, p < .001$. Adding the interaction terms to the regression model explained an additional 1.2% of the variation in posttest scores and this change in R^2 was significant, $F(7,157) = 48.0, p < .001$. When all seven variables were included in stage three of the regression model, only two variables were significant predictors of posttest score: pretest score and the interaction between pretest score and the low-high guidance condition. The pretest score uniquely explained 27% of the variation in posttest score and the interaction between pretest and the low-high guidance condition uniquely explained .8% of the variation in posttest score.

Table 5

Summary of Hierarchical Regression Analysis for Variables predicting Posttest Counting

Variable	β	t	sr^2	R	R^2	ΔR^2
Step 1				0.817	0.667	0.667
Pretest TEN Counting	0.765	18.056*	0.667			
Step 2				0.818	0.670	0.003
Pretest TEN Counting	0.761	17.557*	0.637			
Low-High Condition	0.606	-0.280	0.001			
High-Low Condition	-2.552	-1.178	0.003			
High-High Condition	-1.078	-0.495	0.001			
Step 3				0.826	0.682	0.012
Pretest TEN Counting	0.854	11.597*	0.272			
Low-High Condition	14.541	1.927	0.008			

Table 5 (continued)

High-Low Condition	-0.542	-0.075	0.001
High-High Condition	11.868	1.276	0.003
Low-High*Pretest	-0.240	-2.096*	0.008
High-Low*Pretest	-0.026	-0.236	0.001
High-High*Pretest	-0.192	-1.447	0.004

Note. N=165; * $p < .05$

Figure 3 shows the interaction between counting scores and condition. The graph shows that students with pretest scores below 64.2 (highlighted by the reference line at mean of all pretest scores, $Y = 64.2$) showed the highest posttest scores in the high to low condition, followed by the consistent high condition, the consistent low condition, and then the high to low condition. Students with pretest scores above 64.2 had the highest posttest scores in the consistent low condition, followed by the high to low condition, the consistent high condition, and finally the high to low condition.

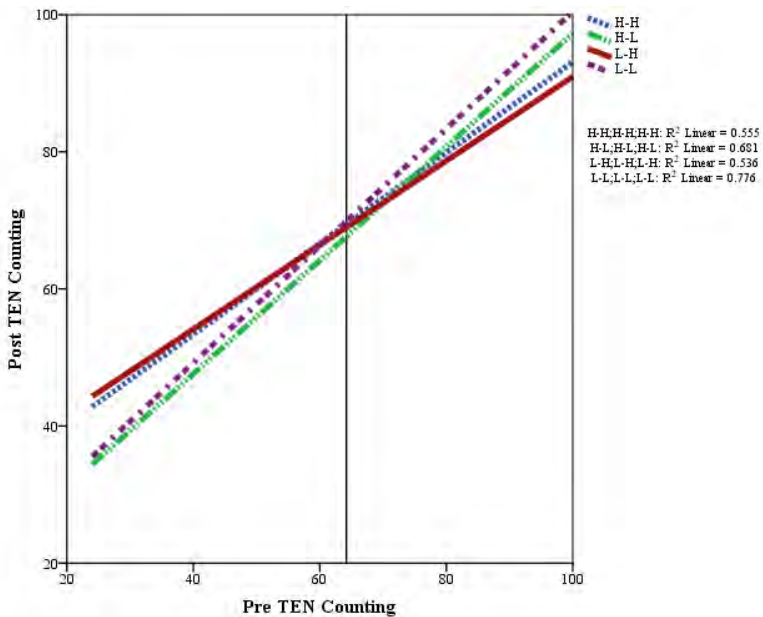


Figure 3. Students' performance on the counting task of the Test of Early Numeracy (TEN) with reference line at mean, $Y = 64.2$.

The same hierarchical linear regression model was performed with delayed TEN counting posttest as the dependent variable but these findings did not remain, only pretest was a significant predictor of delayed posttest score. Subsequent analyses were performed after removing students who scored with the highest score on the counting with manipulatives pretest (five out of five) to account for ceiling effects but these models showed no significant predictors of TEN posttest scores other than TEN pretest scores.

Discussion

Based on the limited prior research on guidance and prior knowledge, we made several hypotheses. We predicted the transitioning high to low and low to high guidance conditions would be the best for students' learning on the intervention content specific task, transfer task, and number sense task. Regarding prior knowledge, we predicted an interaction effect; students with low prior knowledge would perform best on counting with manipulatives with consistent high guidance or high to low guidance on counting with manipulatives. Students with high prior knowledge would perform best with consistent low guidance or low to high guidance when learning with manipulatives. Overall, none of these findings were supported by the results of this study, with most comparisons between conditions showing no difference in learning, even after controlling for prior knowledge.

Overall, student performance on counting with pennies and nickel strips did not differ between conditions even after controlling for pretest scores and possible ceiling effects. These current findings contradict prior research, which has typically shown high guidance groups outperform control groups when highly guided instruction is implemented (e.g., [Carbonneau et al., 2013](#); [Hunt, 2014](#); [Terwel et al., 2009](#)). There are several possible explanations for this finding. First, guidance may not be a moderator of learning for counting to ten with manipulatives. This skill may not require explicit explanation or questioning from an instructor. Simply allowing students to practice and count on their own may be all that is needed. Another explanation could be the ineffectiveness of this task for assessing deeper learning. The questions included in the high guidance modifications targeted deeper learning, as they focused on comparing columns and noticing similarities and differences in the quantities. The task assessing the intervention content only had students count pennies and nickel strips, which did not relate to the questions used in the high

guidance modifications. Perhaps asking questions to target this deeper learning would have shown differences in learning by condition.

Overall, student performance on the transfer task also did not differ across conditions even after controlling for pretest scores and possible ceiling effects. As with the intervention content specific task, it could be guidance is not a moderator of learning for transfer to counting on from five. However, based on the research on guidance, it is surprising this study did not show differences between conditions. Specifically, the meta-analysis by Carbonneau et al. (2013) suggested that conditions that implemented low guidance (i.e., high to low and low to high) would show greater performance on transfer tasks. Typically, allowing students time to practice while also incorporating guidance (i.e., the high to low or low to high guidance conditions) fosters deeper learning. The contradictory findings of this study could indicate an issue with this measure of transfer as we only included two items to yield a score of 0-2. Perhaps a longer or more in depth test of transfer would have provided better insight into students' learning for transfer.

Student performance on the four tasks for the Test of Early Numeracy showed a difference between conditions for the counting task only, where students counted as high as they could, up to 100, for one minute. After controlling for pretest scores, students in the low to high guidance condition scored significantly lower than students in the consistently low guidance condition. The consistently high and high to low guidance groups did not perform significantly different from the consistently low guidance group. These results indicate that providing students with time to practice alone followed by providing guidance can hinder counting fluency; it was better to allow students to practice counting with manipulative on their own with no additional guidance or questions. These results deviate from prior research that included assessments of number sense, where student performance was significantly higher after high guidance instruction compared to low guidance instruction on counting tasks (e.g., Carbonneau and Marley, 2015, found high guidance positively impacted student's conceptual knowledge as measured by magnitude comparison). Perhaps providing guidance in the form of additional questions can distract students from the main task of counting. But, it is interesting students performed lower in the low to high guidance condition but not the high to low condition.

We hypothesized that after controlling for pretest scores we would find interaction effects; students with higher prior knowledge would excel with less guidance while students with lower prior knowledge would excel with more guidance. As we saw no significant results with any other measures we investigated this hypothesis by graphing scores on the TEN counting task. *Figure 3* shows students' pretest and posttest scores on the counting task for each condition. It is clear that students in each condition had a different relationship between pre and posttest performance and this relationship changed dependent upon if students had a pretest score above or below the mean (see reference line $Y = 64.2$). Looking to students in the low to high guidance condition, we see that for students with lower prior knowledge (i.e., below 64.2) their posttest scores were higher than the other conditions. For students with higher prior knowledge (i.e., above the mean of 64.2) their posttest scores were the lowest of all conditions. This graph brings some clarity to the finding from our hierarchical linear regression that students in the low to high condition performed significantly lower than students in the consistent low guidance condition. *Figure 3* shows this clearly for students with higher prior knowledge but not for students with lower prior knowledge. We decided to look at the same hierarchical linear regression but to first analyze students with low prior knowledge (i.e., less than the mean of 64.2) and then high prior knowledge (i.e., greater than the mean of 64.2) but these results revealed no significant predictors of counting scores beyond pretest counting scores.

Based on the overall findings for this study and the tasks used, the only difference in learning was found for the low-high guidance condition, where students performed significantly lower than the consistent low guidance condition on the counting portion of the Test of Early Numeracy. On the TEN counting posttest, students in the low to high guidance condition scored significantly lower when compared to the consistently low guidance condition. However, with this being the only measure to show a difference, we cannot draw conclusions about the optimal level and timing of guidance for learning with mathematics manipulatives.

Conclusions

This research study indicates that level and timing of teacher guidance is not necessary to consider when teaching kindergarten students to count to ten with manipulatives. Controlling for prior knowledge (i.e., pretest) did not impact these results. While not statistically significant, we believe these results are of practical significance to teachers and researchers. These results indicate that the counting tasks already being implemented in the classrooms at this school district are providing enough instruction for this content.

While providing guidance for this task did not appear to impact student learning, these results may not be found for more complex or challenging tasks. The only conclusion we can draw from this study is that for this less complex and age appropriate task, the amount and timing of guidance was not important. Future studies can provide further insight into guidance and its impact on student learning.

Future research

Future research should further advance this research on guidance during instruction as well as mitigate some of the limitations we noted. As discussed earlier, some teachers requested to nominate students who they felt would most benefit from the intervention so as not to pull too many students from class. As such, this reduces our ability to generalize these findings to all students. Another limitation relates to the design of our study. The short duration of the intervention (one week, 6-9 minutes per day) may not be sufficient to impact students' understanding differently between conditions. Further, the measures used in this study were chosen and created to be short, and as such may not have been long or complex enough to determine if deeper learning occurred. Questions that target comparisons between number columns (e.g., "how many more pennies are in the seven column than the five column?") would provide more insight into whether or not deeper, more meaningful learning took place beyond simply counting pennies and nickel strips on a board. We also noted that the lack of statistically significant findings could suggest that the counting tasks already being implemented in the students' classrooms were sufficient for teaching counting from one to ten.

It is also well worth considering the knowledge students are already coming to the classroom with. As such, more in depth pretests could be implemented.

Future research should also focus on variations of the timing and level of guidance with other tasks and age groups. The current research design could also be implemented with preschool students. Preschool students do not have the same base level of knowledge for counting in general and counting with manipulatives specifically. Perhaps implementing this research with preschool students would show differences in learning based on the timing and level of guidance.

Future research could also implement a different research design, such as a single subject design to compare the extent to which guidance is associated with fewer trials to mastery. Given the importance of ensuring fidelity with this research on teacher guidance, single subject design would allow researchers to understand how guidance truly impacts individual students, without the noise created by other students being in the class and impacting student learning. However, given that teaching occurs in classrooms with many students, implementing studies in real classrooms with real teachers is also important to determine the generalizability of these findings.

References

- Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tennenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology, 103*(1), 1-18. doi: 10.1037/a0021017
- Baroody, A. J., Purpura, D. J., Eiland, M. D., & Reid, E. E. (2015). The impact of highly and minimally guided discovery instruction on promoting the learning of reasoning strategies for basic add-1 and doubles combinations. *Early Childhood Research Quarterly, 30*(Part A), 93–105. doi: 10.1016/j.ecresq.2014.09.003
- Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review, 31*(1), 21–32.
- Bruner, J. S. (1966). *Toward a theory of instruction*. Cambridge, Mass., Belknap Press of Harvard University.
- Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Educational Psychology, 105*(2), 380-400. doi: 10.1037/a0031084

- Carbonneau, K., & Marley, S. (2015). Instructional guidance and realism of manipulatives influence preschool children's mathematics learning. *The Journal of Experimental Education*, 1–19. <https://doi.org/10.1080/00220973.2014.989306>
- Chen, O., Kalyuga, S., & Sweller, J. (2015). The worked example effect, the generation effect, and element interactivity. *Journal of Educational Psychology*, 107(3), 689–704. doi: 10.1037/edu0000018
- Clarke, B., & Shinn, M. (2004). *Test of early numeracy (ten): Administration and scoring of aimsweb early numeracy measures for use with aimsweb*. Eden Prairie, MN: Edformation Inc.
- Cobb, P. (1995). Cultural tools and mathematical learning: A case study. *Journal for Research in Mathematics Education*, 26(4), 362-385. DOI: 10.2307/749480
- Darch, C., Carnine, D., & Gersten, R. (1984). Explicit instruction in mathematics problem solving. *Journal of Educational Research*, 77(6), 351.
- DeLoache, J. S. (2000). Dual representation and young children's use of scale models. *Child Development*, 71(2), 329–338. doi:10.1111/1467-8624.00148
- Fisher, K. R., Hirsh-Pasek, K., Newcombe, N., Golinkoff, R. M. (2013). Taking shape: Supporting preschoolers' acquisition of geometric knowledge through guided play. *Child Development*, 84(6), 1872-8. <https://doi.org/10.1111/cdev.12091>
- Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., ... Jancek, D. (2003). Explicitly teaching for transfer: Effects on third-grade students' mathematical problem solving. *Journal of Educational Psychology*, 95(2), 293–305. <https://doi.org/10.1037/0022-0663.95.2.293>
- Fuson, K. C. (2009). Avoiding misinterpretations of Piaget and Vygotsky: Mathematical teaching without learning, learning without teaching, or helpful learning-path teaching? *Cognitive Development*, 24(4), 343–361. <https://doi.org/10.1016/j.cogdev.2009.09.009>
- Fyfe, E. R., Rittle-Johnson, B., & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology*, 104(4), 1094-1108. <http://psycnet.apa.org/doi/10.1037/a0028389>

- Fyfe, E. R., & Rittle-Johnson, B. (2016). The benefits of computer-generated feedback for mathematics problem solving. *Journal of Experimental Child Psychology*, 147, 140–151. <https://doi.org/10.1016/j.jecp.2016.03.009>
- Harcourt Assessment Inc. (2002). *Stanford achievement test, tenth edition*. San Antonio, TX: Harcourt Assessment Inc.
- Horan, E. M. (2017). *Guidance and structure in mathematics instruction: How much guidance do students need? An intervention study on kindergarten mathematics with manipulatives* (Doctoral dissertation). Retrieved from https://galileo-usg-uga-primo.hosted.exlibrisgroup.com:443/UGA:UGA:01GALI_USG_ALMA71186079720002931
- Hunt, J. H. (2014). Effects of a Supplemental Intervention Focused in Equivalency Concepts for Students with Varying Abilities. *Remedial and Special Education*, 35(3), 135–144. <https://doi.org/10.1177%2F0741932513507780>
- Jitendra, A. K., Rodriguez, M., Kanive, R., Huang, J.-P., Church, C., Corroy, K. A., & Zaslofsky, A. (2013). Impact of small-group tutoring interventions on the mathematical problem solving and achievement of third-grade students with mathematics difficulties. *Learning Disability Quarterly*, 36(1), 21-35. doi:10.1177/0731948712457561
- Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, 19(4), 509-539. doi: 10.1007/s10648-007-9054-3
- Kroesbergen, E. H., & van Luit, J. E. H. (2002). Teaching multiplication to low math performers: Guided versus structured instruction. *Instructional Science*, 30(5), 361–378. <https://doi.org/10.1023/A:1019880913714>
- Kroesbergen, E. H., & Van Luit, J. E. H. (2005). Constructivist mathematics education for students with mild mental retardation. *European Journal of Special Needs Education*, 20(1), 107–116. <https://doi.org/10.1080/0885625042000319115>
- Laski, E. V., Jor'dan, J. R., Daoust, C., & Murray, A. K. (2015). What makes mathematics manipulatives effective? Lessons From cognitive science and montessori education. *SAGE Open*, 5(2), 1-8. doi: 10.1177/2158244015589588

- Martinez, R. S., Missall, K. N., Graney, S. B., Aricak, O. T., & Clarke, B. (2009). Technical adequacy of early numeracy curriculum-based measurement in kindergarten. *Assessment for Effective Intervention*, 34(2), 116-125. doi: 10.1177/1534508408326204
- Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, 59(1), 14-19. doi: 10.1037/0003-066X.59.1.14
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics*. Washington D.C.: National Governors Association Center for Best Practices, Council of Chief State School Officers. Retrieved from <http://www.corestandards.org/the-standards>
- National Research Council. (2009). *Mathematics Learning in Early Childhood: Paths toward Excellence and Equity* (No. 978-0-3091-2806-3). Washington, D.C.
- Piaget, J. (1977). *Epistemology and psychology of functions*. Boston, MA: D. Reidel Publishing Co.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York, NY, US: Oxford University Press.
- Salvia, J., & Ysseldyke, J. E. (2001). *Assessment (8th ed.)*. Boston, MA: Houghton Mifflin.
- Schwartz, J. E. (1992). "Silent teacher" and mathematics as reasoning. *Arithmetic Teacher*, 40(2), 122-124.
- Sengupta-Irving, T. & Enyedy, N. (2014). Why engaging in mathematical practices may explain stronger outcomes in affect and engagement: Comparing student-driven with highly guided inquiry. *Journal of the Learning Sciences*, 24(4), 550-592. <https://doi-org.proxy-remote.galib.uga.edu/10.1080/10508406.2014.928214>
- Sowell, E. J. (1989). Effects of manipulative materials in mathematics instruction. *Journal for Research in Mathematics Education*, 20(5), 498-505. doi:10.2307/749423
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York, NY : Springer.
- Terwel, J., van Oers, B., van Dijk, I., & van den Eeden, P. (2009). Are representations to be provided or generated in primary mathematics

- education? Effects on transfer. *Educational Research and Evaluation*, 15(1), 25–44. <https://doi.org/10.1080/13803610802481265>
- Tournaki, N. (2003). The differential effects of teaching addition through strategy instruction versus drill and practice to students with and without learning disabilities. *Journal of Learning Disabilities*, 36(5), 449-458. <https://doi.org/10.1177%2F00222194030360050601>
- Uttal, D. H., O’Doherty, K., Newland, R., Hand, L. L., & DeLoache, J. (2009). Dual representation and the linking of concrete and symbolic representations. *Child Development Perspectives*, 3(3), 156–159. [doi:10.1111/j.1750-8606.2009.00097.x](https://doi.org/10.1111/j.1750-8606.2009.00097.x)
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Woodcock, R., & Johnson, M. (1989). *Woodcock-Johnson Psycho-Educational Battery-Revised*. DLM Teaching Resources.

Erin Horan, PhD, is a Postdoctoral Research Analyst in the Office of Quality, Curriculum, and Assessment at the Association of American Colleges and Universities (AAC&U).

Martha Carr, PhD, was a University of Georgia (UGA) Aderhold Distinguished Professor of Educational Psychology and Research Fellow of the UGA Institute of Behavioral Research.

Contact Address: Erin Horan, Association of American Colleges and Universities, 1818 R Street NW, Washington DC, USA. Email: ehoran00@gmail.com